ERIC Document Resume	See Instruction Sheet	DATE OF RESUME	DATE RECEIVED

A. PUBLICATION DATA

1. ERIC NUMBER	2. ORIGINATOR CONTROL NO(S).	3. IDENTIFICATION NO.	DO NOT WRITE IN SPACE BELOW
ED002239	N.A.	1193	

4. AUTHOR(S): Alter, Millicent, and others

5. TITLE OF DOCUMENT ABSTRACTED: The Effectiveness of Confirmation Plus Trinket Reinforcers in Young Children

6. PUBLICATION TITLE: N.A.

7. EDITOR(S): N.A.

8. PLACE OF ORIGIN: The Center for Programed Instruction, 365 West End Avenue, New York, New York

9. PUBLISHER OR SOURCE: The Center for Programed Instruction, 365 West End Avenue, New York New York

CITY, STATE

TELE. EXT.

INST/LOC CODE

10. VOL.	11. PAGES	12. MO./YR.	13. TOTAL PP.
N.A.	N.A.	N.A.	23

MICROFILM COPY AVAILABLE: YES ☒ NO ☐

COPYRIGHT GRANTED: YES ☐ NO ☒

COPYRIGHT RELEASE FILED: YES ☐ NO ☐

TYPE OF REPORT: PROPOSAL ☐ IN-PROGRESS ☐ FINAL ☒ FOLLOW-UP ☐

OTHER ERIC REFERENCES:

ERIC SATELLITE CODE

B. INDEXING DATA

14. SUBJECT AREA

15. TYPE OF ENDEAVOR: Report

16. INDEXING TERMS (6-12)

reinforcers		motivation techniques		
reinforcement		preschool		
programed learning		rewards		
instructional technology		kindergarten		

17. ABSTRACT (MAX. 250 WORDS)

A designed program for preschool and kindergarten children who have not yet learned to read and write was described. It dealt with number concepts and numerals and used a multiple choice pictorial format. In the typical programed instruction situation, the student read a frame and made a response. When he made the correct response, he was informed that he was right. "Being right" may be thought of as a secondary reinforcer. For young children who have not yet become "socialized" to the extent that older students have, "being right" might not serve as a potent or even effective reinforcer. In view of this, a comparison of the effectiveness of confirmation alone with confirmation plus the receipt of trinkets which are likely to appeal to young children was made.

It was hypothesized that, in five- and six-year-old children, confirmation would not be as effective to reinforce as confirmation plus trinkets. Two experimental groups were formed, both of which worked through a preliterate program designed to teach number concepts and numerals. One received simple confirmation of correct responses, while the other received confirmation plus reinforcement with trinkets. No significant differences were found between the two groups.

ERIC No.

18. SOURCE OF ABSTRACT

C. FUNDING DATA

19. SUPPORTING ORGANIZATION	20. DIVISION OR DEPT.	21. CONTRACT NO. / PROJECT NO.
	22. TOTAL COST OF PROJECT	
	23. DURATION OF PROJECT	U.S. Department of Health, Education, and Welfare Office of Education

ERIC Form 1

1193

T 18

ED002239

INSTRUCTIONAL TECHNOLOGY - U.S.C.

The Effectiveness of Confirmation Plus Trinket Reinforcers in Young Children

by

Millicent Alter, Lewis Eigen and Shirley King

The Center for Programed Instruction, Inc.

REPRINTED BY PERMISSION
U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN PRINTED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Published & Distributed by

The Center for Programed Instruction
365 West End Ave., New York 24, N. Y.

A non-Profit Educational Organization

All proceeds from sale of articles are used to defray printing costs and to further research in Programed Instruction

0358

The Effectiveness of Confirmation Plus Trinket Reinforcers in Young Children[1]

Millicent Alter, Lewis Eigen and Shirley King

The Center for Programed Instruction

A preliterate program is one which does not assume reading or writing ability on the part of the student. A large portion of the nonliterate population is composed of young children who have not yet been taught to read and write, hence the term "preliterate." A program which does not assume literacy is necessarily more limited in its techniques than one which does in that (a) information cannot be presented in the form of the written word, and (b) the student cannot be required to write words in response. In a program addressed to preliterates, most of whom are quite young, certain other restrictions arise regarding the degree of coordination required to make the response, and the nature of the pictures and words used to convey the material.

In programming scientific concepts for first graders Kieslar (1961) used a multiple choice pictorial format accompanied by a recorded commentary. The program described below, dealing with number concepts and numerals, was intended for preschool and kindergarten children. It also used a multiple choice pictorial format, but without an accompanying commentary. The studies presented here were conducted during the course of a trial run of this program.

In the typical programmed instruction situation the student reads a frame, and makes a response. Hopefully, the student's behavior has been so effectively shaped that the probability of his making a correct response is high. When he emits the correct response he is informed that he is right. It is generally assumed that this reinforces the behavior we are trying to shape.

"Being right" may be thought of as a secondary reinforcer: an originally neutral state of affairs which has acquired reinforcing properties through the behavior of a society which makes the receipt of other reinforcers contingent upon "being right." The experimenters felt that for young children, who have not yet become "socialized" to the extent that older students have, "being right" might not serve as a potent, or even effective, reinforcer. In view of this, the experimenters felt it would be of interest to compare the effectiveness of confirmation alone with confirmation plus the receipt of trinkets which are likely to appeal to young children.

[1]This study was made possible through the financial support of the Fund for the Advancement of Education of the Ford Foundation.

Experiment 1

Program

The program[2] consisted of 262 frames designed to teach the quantitative concepts "oneness" through "nineness." The term "fourness," for example, is used to indicate that which all sets of four objects have in common quantitatively regardless of their qualitative dissimilarities. The program was designed also to teach the symbolic representation for each quantity, the numerals 1, 2, 3, 4, etc., but did not attempt in any way to teach counting. The program was arranged in nine sequences of frames, each sequence covering first the quantitative concept and then the numeral in order from one to nine.

Two sequences of frames are shown in the Appendix, the first dealing with the concept "oneness," the second dealing with the numeral '4'. All frames in the program had the same format. This consisted of a top portion containing a "question" or "stimulus" picture, and a lower portion containing three alternative "answer" or "response" pictures. The subject was instructed to select the alternative which had the same number or amount in it as the top picture had. He was required to make three types of matches: (a) number of objects to number of objects (see Frame 1 of Appendix), (b) numeral to number of objects (see Frame 105), and (c) number of objects to numeral (see Frame 112).

Two cueing techniques were used:

(a) Identity of objects: The correct response alternative depicted the same sort of objects (dogs, balls) as the stimulus picture, while the incorrect alternatives depicted different objects. This is illustrated in Frames 1 and 2 of the Appendix.

(b) Color: The correct response alternative was the same color as the stimulus picture while the incorrect alternatives were a different color. This is illustrated in Frame 105 of the Appendix.

Measurements

1. Time: The time taken by each subject to complete the program was recorded.

2. Errors: Errors made by each subject on the program were recorded so as to indicate the frame and the particular erroneous response chosen. These were summed to yield the total number of errors made by each subject.

3. Pre- and Posttest: Each subject was tested on the programmed material before and after going through the program. The same test was used for pre- and posttesting. This consisted of a 21 frame booklet which had the same format as the program, and covered the same material, the quantitative concepts and numerals one through nine. The test required the same types of matches as the program. The test frames differed from the program frames in that the former had all cues withdrawn, and used pictures which did not appear in the program.

Subjects

The subjects were 16 five to six year old children (8 male and 8 female).

[2]Written by Shirley King and Gloria Johns of the Dalton School, New York City.

They were drawn from three different classes at the Dalton School, a private school in New York City. Although the socioeconomic background of the subjects is not known in detail, it is known that their backgrounds were, in general, at least middle class.

Procedure

Subjects went through the program and the pre-and posttests in unused classrooms in their own school building. Both the program and the tests were administered personally by the experimenters who sat with the subjects and showed them each frame. No time limits were set for either the tests or the program.

Each subject was pretested with the above-mentioned test. He was asked to choose the bottom picture which had the same number in it as the top picture had. When he made his selection, the subject was not told whether he was right or wrong, but simply shown the next frame.

Subjects getting 18 or more of the 21 items correct on the pretest were dropped from this study. Of the remaining 20 subjects 16 were assigned to two experimental groups with two considerations: (a) within each class the number of subjects in each experimental group was approximately equal, and (b) the two experimental groups were matched as closely as possible in mean and variance of the pretest scores.

The two experimental groups are referred to as "trinket" and "regular." "Regular" subjects were told on each frame whether they had chosen correctly or incorrectly. If the latter, they were shown which alternative was correct before advancing to the next frame. They could have been required to continue choosing until they had selected the correct alternative before advancing to the next frame. The former procedure was chosen both because it is analogous to the typical programmed instructional situation and because it better fitted the requirements of the study. The "trinket" group proceeded in the same manner and, in addition, was given small prizes (foil stars, marbles, and balloons) on some of the occasions on which they were correct. Each of these subjects was given trinkets for each of the first five correct responses. After this he was put on a variable ratio schedule and given trinkets for approximately 25% of his correct responses.

The experimental room contained the child serving as the subject and two experimenters. One experimenter showed the program to the child frame-by-frame and told him if he was correct, and if not which alternative was correct. The other administered the trinkets when called for, recorded the errors made by the subject and noted the time he took.

The program was split into three segments which were presented in three experimental sessions. Each session lasted for an average of 15 minutes, although they ranged from a minimum of 7 to a maximum of 31 minutes. Subjects generally went through the program on three consecutive weekdays and were tested on the fourth day. Three of the subjects were absent for one day, and so continued where they had left off in the program when they returned. Although the subjects did not manifest strong physical resistance when taken from their home rooms to participate in the experiment, they could not be described as enthusiastic. This may be due to the nature of the program, but their comments suggested that, in part, it was due to the length of the sessions which is probably excessive for children of this age.

Results and Discussion

Estimates of internal-consistency reliability (Kuder-Richardson #20) on the pre- and posttests were .68 and .89 respectively. The distribution of pretest scores was more or less symmetrical about the mean. The distribution of posttest scores was asymmetrical with a large number of cases piling up at the ceiling end of the scale. Four of the sixteen subjects achieved perfect scores confirming that the test had too low a ceiling.

The median pretest score of the "trinket" and "regular" groups combined was 9.8 (out of a possible 21) correct while the median posttest score was 18.5, a gain of approximately 9 points. This difference was found to be significant beyond the .01 level using a Wilcoxon matched-pairs signed-ranks test. This would seem to indicate that learning took place; however, as will be discussed below, this effect may be specious.

Table 1 compares the "trinket" and "regular" groups on the four measures

Insert Table 1 about here

taken. Mann-Whitney U tests indicated that the groups did not differ significantly from one another in either pre- or posttest score. The groups were compared with respect to number of errors made by the subjects on the program (out of a possible 262) and time taken to complete the program. t's of .15 and 1.46 respectively indicated that the groups did not differ significantly in these respects either.

Thus, we cannot demonstrate that the "trinket" and "regular" conditions are not equally effective in teaching this material. These findings may be interpreted in at least two ways: either the trinkets were not reinforcing to the children (this was not checked independently), or, they were reinforcing to the children, but the combination of trinkets plus confirmation was no more reinforcing than confirmation alone. That is, the trinkets might have supplanted confirmation which, in their presence, may have been rendered trivial as a reinforcer. A third explanation will be discussed in Experiment 2 which attempted to examine what was taught by the program.

Table 2 presents the intercorrelations among the measures taken. Pearson

Insert Table 2 about here

r's were used. Due to the fact that the posttest had too low a ceiling, their use was not strictly appropriate and the coefficients must be regarded as attenuated to some degree. They did not, however, disagree radically with Spearman rhos which were also computed. The only significant ($p < .01$) correlation was that between the posttest score and the number of errors made by the subject on the program (- .79) which indicated that the greater the number of errors made on the program the poorer the posttest performance. The low (+.07) correlation between the pre- and posttests suggested that perhaps pretest score was not an ideal matching variable.

It is of interest to examine the subjects' program and posttest performance with respect to each of the sequences (concept and numeral combined), one through nine. Figure 1 combines the data of "trinket" and "regular" subjects to yield

Insert Figure 1 about here

error rates[3] on both the program and the posttest as a function of the numerical sequence. As can be seen, the error rate on the program declined between one and three, remained constant from three to five, and showed a consistent increase between six and nine. Analogously, on the posttest the error rate was approximately constant from one to six, and increased between six and nine. Since the format of the test dictated that a subject would, by chance alone, choose the correct answer one third of the time it appeared that subjects were doing no better than chance on the test sequence dealing with the concept and numeral nine. There was one particularly easy item testing the concept "eightness" which yielded an extremely low error rate. The error rate on eight is shown both with and without (dotted line) this item. The significance of these error rates will be discussed following a description of Experiment 2.

Experiment 2

It was noted during the administration of the pretest that many, if not all, of the subjects seemed to be somewhat unclear as to what was expected of them. Although they understood that they were required to point to one of the three bottom pictures before they could proceed to the next frame, they did not seem to grasp the principle of making a comparison between the stimulus picture and the three response alternatives.

From their experiences in the first few administrations of the program the experimenters gained the impression that the children were learning the behavior required in test- and program-taking, e.g. making comparisons. If this was what was learned, this alone could have accounted for part of the expected pre- to posttest gain. This suggested that the study to be described might provide a better baseline against which to evaluate the effectiveness of the program in teaching quantitative concepts and numerals.

Subjects

Four children (2 male and 2 female) drawn from the subject pool described in Experiment 1.

Program and Measurements

The same as described in Experiment 1. Time to complete the program and data relating to error rates on the program are not reported for this study.

Procedure

This study and Experiment 1 were run concurrently. Subjects were pretested and selected in the manner described in Experiment 1. The Subjects in Experiment 2, however, worked through only one-third of the program (87 of the total 262 frames) which covered the concepts and numerals one through three.

[3] All error rates reported in this paper were computed by dividing the total number of errors made by the total number of errors possible, which, in this case, was equal to the number of subjects x the number of frames.

The decision to use this portion of the program was based upon the subjective observations of the experimenters. It appeared that by the time the subjects had completed the first third of the program they understood what was required of them. If it can be assumed that the decline in the program error rate (Figure 1) between one and three is due to an increasing comprehension of the instructions, the leveling off of the error rate between the numbers three and five indicates that an optimal choice was made.

In all respects other than those specified this procedure was identical to the procedure in Experiment 1. Two of the subjects were reinforced; the other two were nonreinforced. This portion of the program required only one experimental day to administer. Subjects were tested on the following day with the complete posttest covering the sequences (concepts and numerals) one through nine. Thus, these students took only one-third of the program, but were tested on material covered in the entire program.

Results and Discussion

The small sample size in this study precludes the drawing of any definitive conclusions. Table 3 presents the median pre- and posttest scores for both the

Insert Table 3 about here

group that went through the entire program (all subjects in Experiment 1), and the group that went through the partial program (first third only). The median score for the entire-program group rose from 9.8 correct on the pretest to 18.5 on the posttest. As mentioned previously, this was statistically significant at the .01 level. The median score of the partial-program group rose from 9.0 to 18.0, a gain which was of the same order as that of the entire-program group.

It would be more precise to consider only that portion of the programed material to which the partial-program group was not exposed, the sequences (concepts and numerals) four through nine. Table 2 shows the median pre- and posttest scores of the entire- and partial-program groups on those 12 (of the total 21) items which dealt with this material. This data indicates that the group which went through the entire program gained 4.6 points on this portion of the test, whereas the partial-program group gained 4 points.

The fact that the group which had gone through only the first third of the program improved their performance on a test of material covered by the latter two-thirds of the program indicates that they learned something about test-taking behavior during the pre- to posttest interval, or that they learned something about this material outside the experimental situation, or both. Learning outside the experimental situation could be expected to affect the entire- and partial-program groups equally. If the program had taught anything about the number concepts and numerals we would expect that on a test covering the latter two-thirds of the program the group which had gone through this material would perform better than the group which had not.

Mann-Whitney U tests were used to compare the entire- and partial-program groups on their pre- and posttest scores involving the sequences four through nine. There was no statistically significant difference between the groups on either pre- or posttest. Apparently, then, we cannot demonstrate that the entire program is any more effective that the first third (which takes one-fourth the time) in teaching

whatever is taught by this program.

If we choose to give weight to the data of Experiment 2 we might interpret it to mean that (a) there is a remarkable amount of transfer from the material taught in the sequences one through three to a test on the sequences four through nine, or (b) the subjects already knew something about this material and the program taught them how to demonstrate this on a test. Since whatever was taught by the entire program appeared to be taught equally well by the first third, the effectiveness of trinkets as reinforcers may have been washed out by the overlearning which took place on the remaining two-thirds. The small size of the partial-program group precluded a valid comparison of the reinforced and nonreinforced subjects.

The error rate on this program was 23%. This is relatively high with respect to generally accepted standards in programmed instruction. Certain types of frames contributed considerably more to this error rate than others.

The programmers categorized each frame on the basis of certain aspects of their contents. The category of each frame in the Appendix is indicated there. An inspection of the frames reveals the bases for the categorization. In frame type 1, for example, the three response alternatives depict three different kinds of objects. The correct alternative depicts the same sort of object as the stimulus picture. In frame type 2, both the stimulus and all three response alternatives depict the same sort of objects.

The highest error rates occurred on frame types 6 (38%) and 7 (40%). In both of these the correct response depicts objects which differ from those depicted by the stimulus, although one (type 7) or both (type 6) of the erroneous responses depict objects which are similar to the stimulus. Other notably high error rates were generated by frames of type 3 (32%) and type 5 (31%). Type 1 frames, understandably, had the lowest error rate (11%). The remainder of the frames ranged between these extremes.

The error rate by frame is presented in Table 4. There are several points

Insert Table 4 about here

to be made in connection with this. It can be seen that there was a tendency for errors to cluster in that subjects who erred on a given frame all tended to choose the same alternative, (e.g. frames 7, 8, 9, 11, etc.). The more marked clusters fell into one of two groups: alternatives which depicted the same sort of objects that the stimulus depicted (frames 15, 16, 17 for example), and alternatives which were located in the same position (left, right, center) as the correct alternative on the immediately preceding frame (frames 7, 8, 9, 11, etc.). The frames dealing with numerals did not seem to present as much difficulty as those dealing with concepts and yielded somewhat lower error rates.

Evidence, both observational and statistical, suggested that the effectiveness of this program is limited in several ways which might be remedied.

1. Both the test and the program required behavior with which the subjects were unfamiliar, e.g. making a comparison between the stimulus and the response alternatives and matching the two. In such a situation it might be wise to administer a pre-program prior to the pretest, involving similar behavior but different content (e.g. color matching) so that (a) the baseline will not be artificially depressed by the subject's lack of testwiseness, and (b) the efficacy of the program will not be impaired.

During the course of the pretest it was frequently observed that the subjects did not look at the stimulus picture before selecting one of the responses. Holland[4] has suggested a teaching device which, for each frame, requires the subject to press a transparent panel which covers the stimulus picture in order to "set up" the machine after which he may choose one of the alternatives. This requirement might increase the probability that the subject would look at the stimulus panel.

2. Each frame of the program gives the student three alternatives to choose from. By chance alone the subject will choose correctly one-third of the time whether or not he is attending to what is being taught. If "being right" is reinforcing, we are reinforcing undesirable behavior on a highly effective schedule. One obvious solution is to increase the number of alternatives. Realizing that when the number of alternatives is kept within reasonable limits some answers are bound to be correct by chance alone, Holland has suggested a backtracking procedure to ameliorate the situation. When a subject errs on a frame, Holland suggests that he then be required to repeat the preceding frame, and, if he errs on that, to be backtracked to the frame preceding that one. Thus, each frame must be responded to without error at least once.

3. The cueing techniques, rather than facilitating learning, may have impeded it: To the experimenters, the subjects often seemed to be learning and responding to the cues rather than to that which the programmers wanted to teach. In the case of the "identity of objects" cue a frame-by-frame error analysis substantiated this. At the beginning of each sequence relating to quantitative concept correct responses were cued by depicting objects similar to those in the stimulus picture. Subjects showed a marked tendency to continue making matches in terms of the qualitative aspects of the pictures. Frames which required the subject to choose alternatives which were quantitatively similar to the stimulus picture when qualitatively similar pictures were available as erroneous alternatives yielded relatively high error rates.

This phenomenon cannot be statistically demonstrated for the color cue, but then, there was no situation in which this cue conflicted with the correct answer. There is, on the other hand, no evidence that it facilitated learning.

It is difficult to tell, in such a situation, whether the subject is responding to that which the programmer wishes to teach him, or to the cue aspect of the frame alone. In this program, it is only when we extinguish responses to the cues, which are irrelevant to what we are teaching, that we force the subject to attend to the concepts or numerals involved. This seems inefficient since we are teaching irrelevant responses by leading a subject to make the correct response for the wrong reasons, and reinforcing this behavior. Thus, we are, initially at least, bringing the subjects' responses under the control of the wrong stimuli. Whether or not the subject learns "incidentally" is a moot point.

Although the applicability of some of the above comments is restricted to this program, the underlying approach applies to cueing techniques in general.

If, as was suggested by Experiment 2, the latter portion of the program was relatively ineffective, in terms of teaching number concepts and numerals, the

[4]Personal communication. November, 1961.

points mentioned above might account for this. There is, however, another interesting hypothesis which emerges from a comparison of the data in Figure 1 with data gathered in span of apprehension studies.

It was felt originally by the programmers that it would be possible to teach number concepts and numerals independently of counting behavior. Most of the subjects in this study either could not count accurately or could not count at all, and the program made no attempt to teach this skill.

Span of apprehension studies are concerned with the ability of subjects to estimate accurately the number of objects in a display. In a typical span of apprehension study a subject is presented with a visual display containing a number of black dots on a white background and is asked to make an estimate of the number of dots displayed.

The data in this area, summarized by Woodworth (1956), indicate that a dual process is involved in making accurate estimates: a form of "direct perception" called subitizing which is operative when the number of dots is equal to or less than six, and counting which is operative when the number of dots is greater than six. In situations in which counting is precluded by using short exposure times (of the order of 200 milliseconds) data show that estimates are highly accurate up to six dots, beyond which they become increasingly inaccurate, and increasingly variable. Minturn and Reese (1951) found that if, immediately after each judgment, they informed the subjects of the true number of dots, they decreased the variability of the inaccurate judgments but did not obliterate the trend toward inaccuracy in estimates of more than six dots.

Relating this data to the present study, it will be recalled that the subjects in this study had limited counting ability. If we examine Figure 1, we find that the subjects' inaccuracy increased when they dealt with numbers greater than six both in the program, and in the posttest. There may be many reasons which account for this, but it is interesting to note how well this data coincides with span of apprehension data. If this is the same sort of phenomenon as is involved in span of apprehension studies, then the ability of subjects to learn from that portion of the program dealing with concepts and numerals six through nine is limited unless the subjects are taught to count. Unfortunately, we have no accurate preprogram estimate of the subjects' performance on the numbers six through nine with which to compare his posttest performance, since it was felt that the pretest scores reflected largely the subjects' ability to understand instructions.

Summary

It was hypothesized that, in five and six year old children, confirmation would not be as effective a reinforcer as confirmation plus trinkets. Two experimental groups were formed, both of which worked through a preliterate program designed to teach number concepts and numerals, one receiving simple confirmation of correct responses, while the other received confirmation plus reinforcement with trinkets. No significant differences were found between the two groups. Data from a substudy suggest that the program might not be teaching what it was designed to teach. The cueing techniques used and the concept underlying the program were discussed critically. Similarities between this data and span of apprehension data were noted.

References

Kieslar, E., & McNeil, J. D. Teaching scientific theory to first grade pupils by auto-instructional device. Harvard Educ. Rev., 1961, 31 (1), 73 - 83.

Minturn, A. L. & Reese, T. W. The effect of differential reinforcement on the discrimination of visual number. J. Psychol., 1951, 31, 201 - 231.

Woodworth, R. S. & Schlosberg, H. Experimental Psychology. New York: Holt, 1956.

Table 1

Comparison of "Trinket" and "Regular" Groups

Measure		Experimental Groups Trinket (N = 9)	Regular (N = 7)	p
Pretest	Median	9.9	9.0	$>.05^{b}$
(No. Correct)	Q[a]	2.3	7.5	
Posttest	Median	18.8	18.2	$>.05^{b}$
(No. Correct)	Q	1.8	1.7	
Errors	Mean	62.6	59.0	$>.05^{c}$
	SD	55.9	30.8	
Time	Mean	44.4	36.7	$>.05^{c}$
	SD	14.2	6.3	

[a]Refers to semi-interquartile range.

[b]Result of Mann-Whitney U test.

[c]Result of *t* test.

Table 2

Intercorrelations Among Measures Taken in Experiment 1

(Based on "trinket" and "regular" groups combined)

	Posttest	Time	Program Errors
Pretest	+.07	.00	− .39
Posttest		.00	− .79**
Time			.00

**Significant at .01 level.

Note - All coefficients in this table are Pearson correlation coefficients.

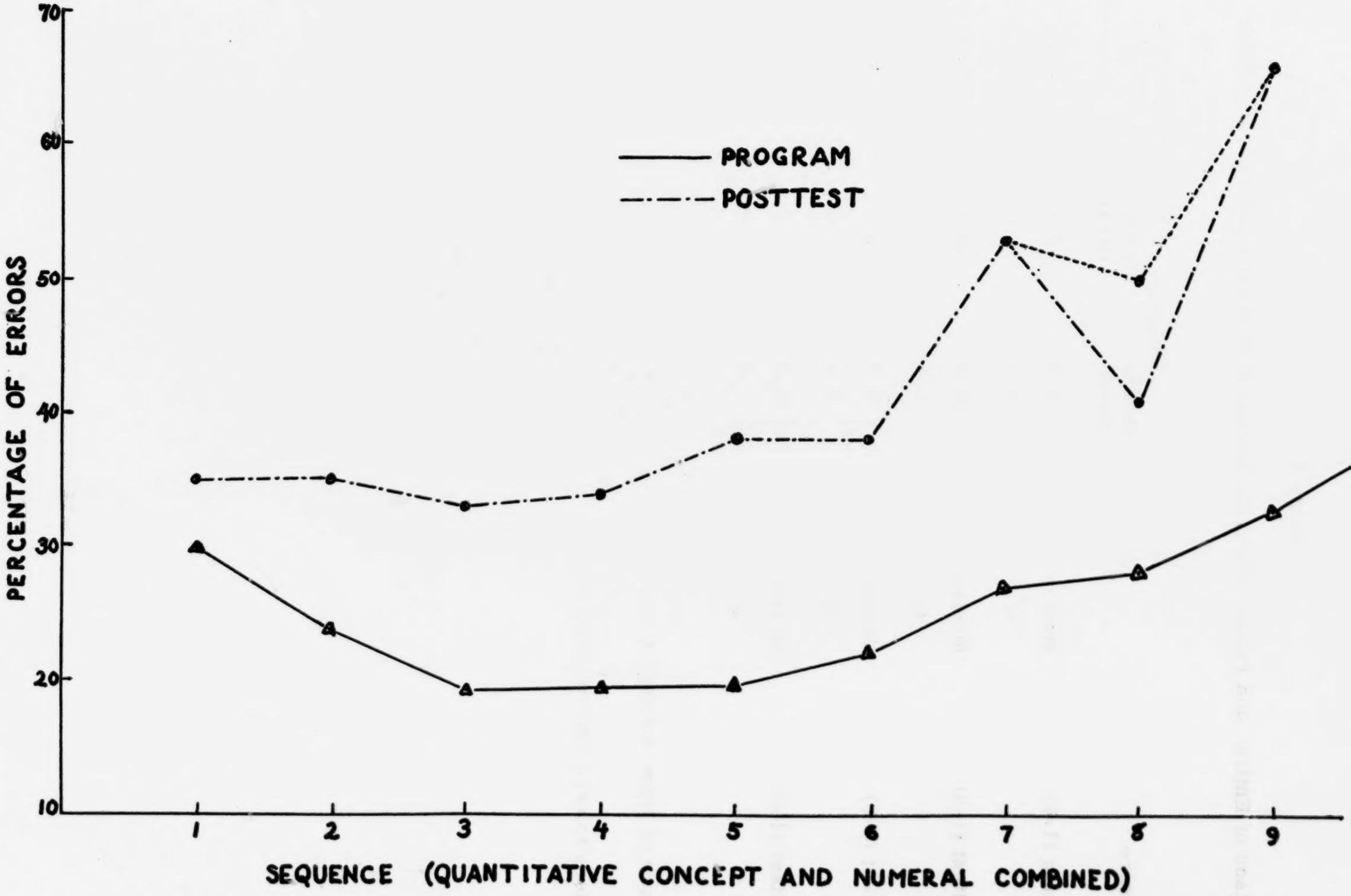


Fig. 1. Error rate on program and posttest as a function of numerical sequence taught or tested. The data are based on all subjects in Experiment 1.

Table 3

Comparison of Entire- and Partial-Program Groups on Complete and Partial Posttest

Measure		Program		p[a]
		Entire (N=16)	Partial (N=4)	
Pretest (1-9)	Median	9.8	9.0	$>.05$
	Q[b]	2.8	1.5	
Posttest (1-9)	Median	18.5	18.0	$>.05$
	Q	1.8	3.5	
Pretest (4-9)	Median	5.2	5.0	$>.05$
	Q	1.4	1.0	
Posttest (4-9)	Median	9.8	9.0	$>.05$
	Q	1.4	2.0	

[a]Result of Mann-Whitney U test.

[b]Refers to semi-interquartile range.

Table 4

Number and Distribution of Errors or Each Frame in Appendix

Frame Number	Type	Alternative			Total Errors
		Left	Center	Right	
1	1	X		1	1
2	1	1	1	X	2
3	1		X		0
4	2		X		0
5	2	X	3		3
6	3	X			0
7	3	8	X	2	10
8	3	X	5	1	6
9	3	7	3	X	10
10	4	3	4	X	7
11	4	2	X	6	8
12	4	1	X	2	3
13	5	3	X	2	5
14	6	X	12		12
15	6	2	5	X	7
16	7	1	5	X	6
17	7	7	X	1	8
18	2		X	1	1
19	3	3		X	3
105	8		X		0
106	8		2	X	2
107	8	X			0
108	9	X	1	1	2
109	10	2	X		2
110	14	1	X	1	2
111	10		3	X	3
112	11		3	X	3
113	12	1	X	2	3
114	13		X	1	1
115	13	X	1	2	3
116	12	X	1		1

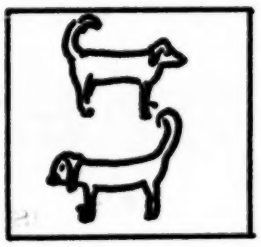
Note - X indicates correct alternative.

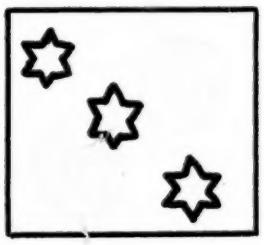
Appendix

Two sequences of frames excerpted from the program. Frames 1 - 20 deal with the concept "oneness." Frames 105 - 116 deal with the numeral '4'. The frame type is shown in parentheses next to the frame number.

These frames were originally drawn in purple. In color-cued frames the stimulus and correct response were drawn in red. The red pictures are grey in this reproduction. The frames in this Appendix have been reduced to approximately one-third their original size.

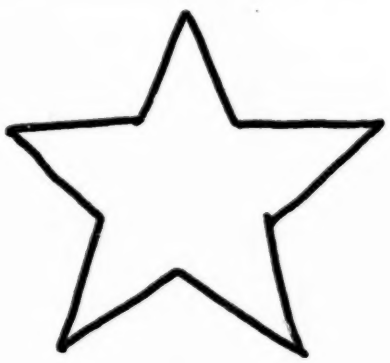
1(1)

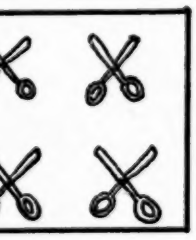


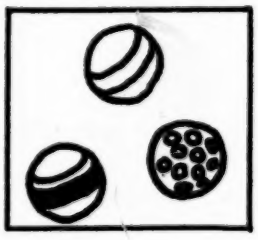


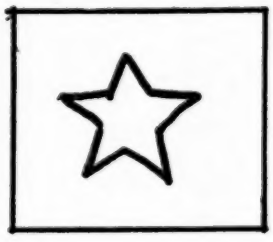
4(2)

2(1)

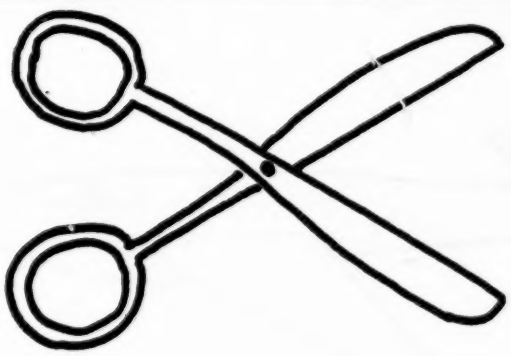


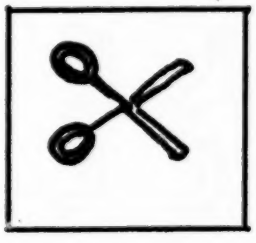


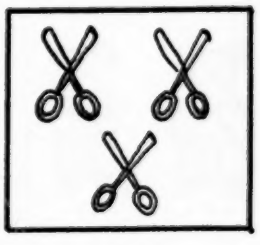


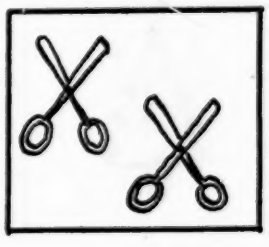


5(2)

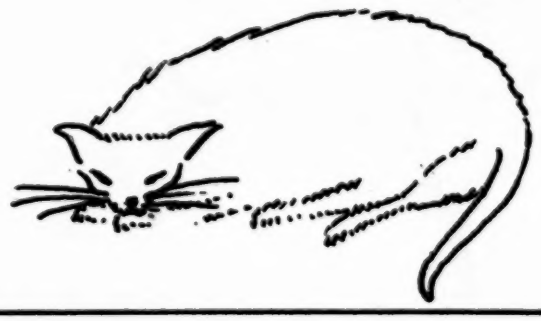




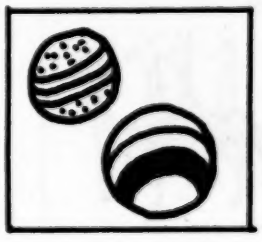




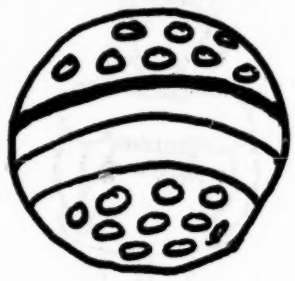
3(1)

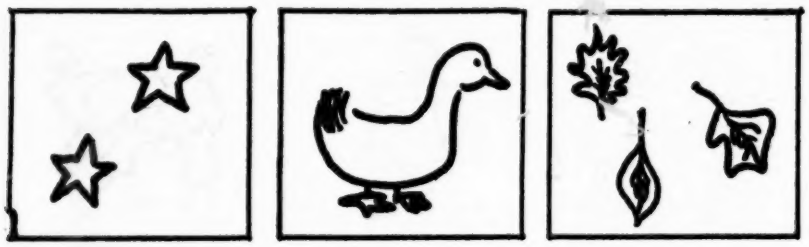


6(3)

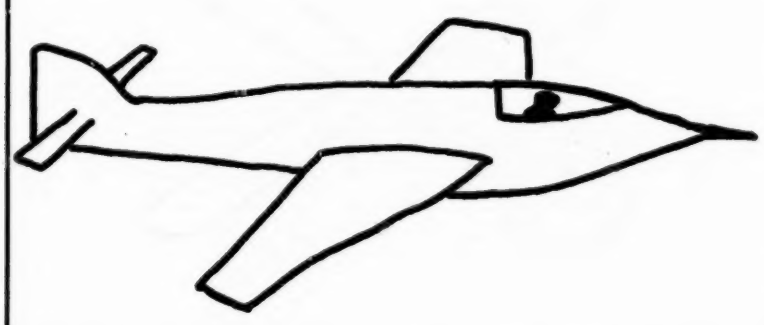


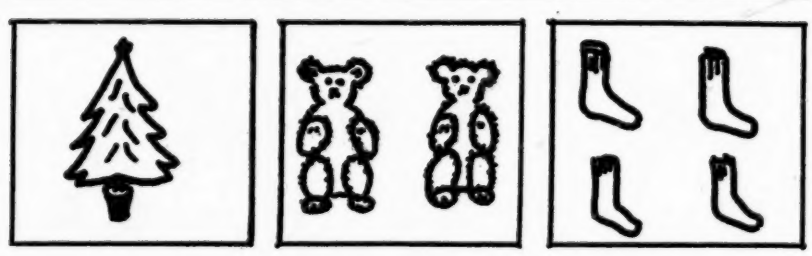
7(3)



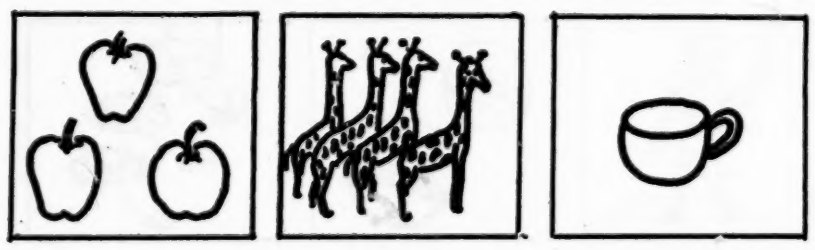


8(3)

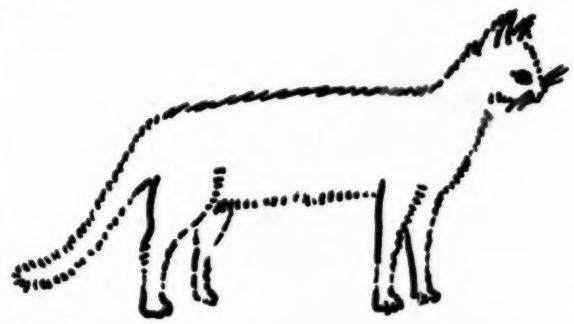


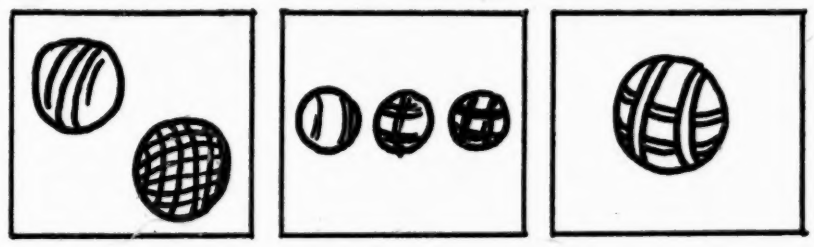


9(3)

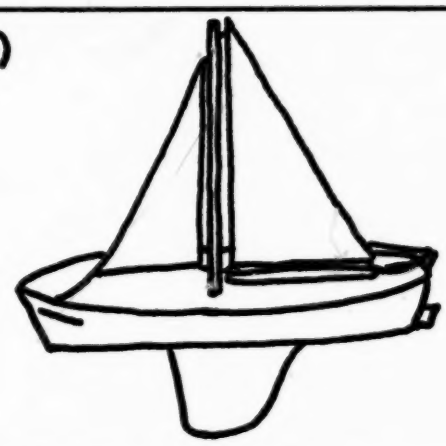


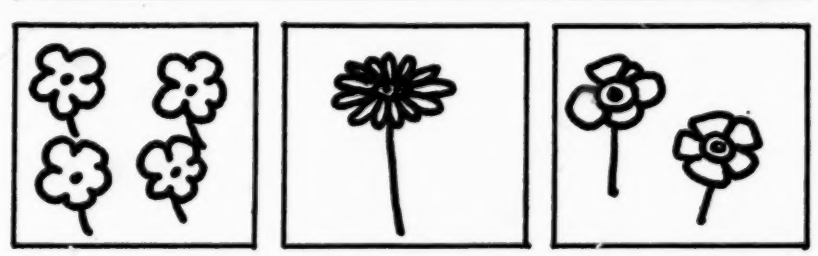
10(4)



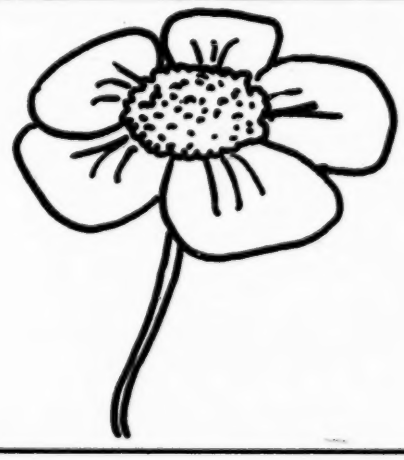


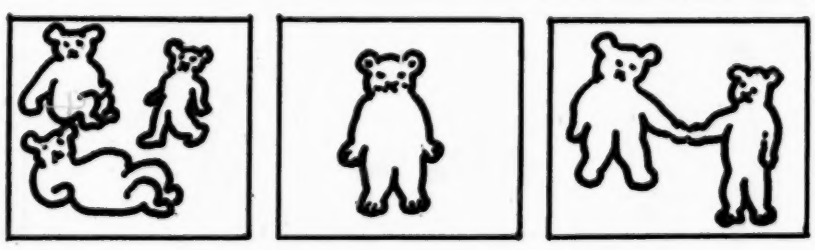
11(4)



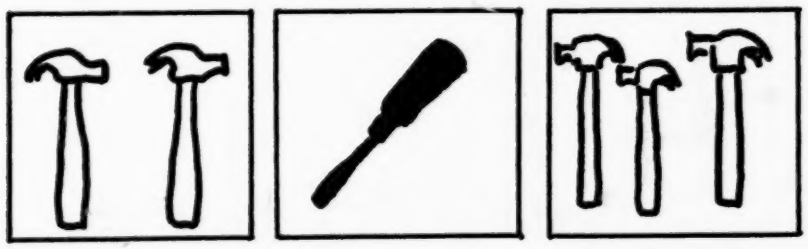


12(4)

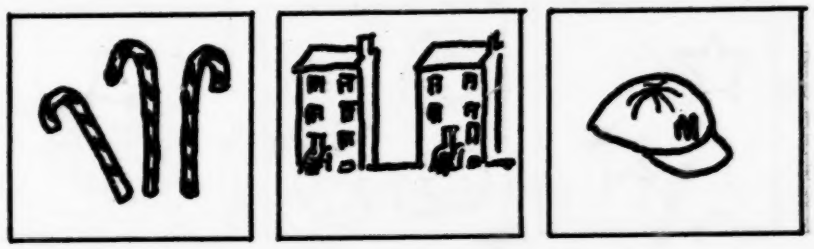




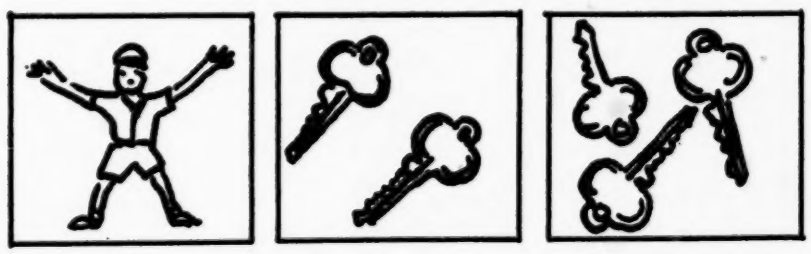
13(5)



16(7)

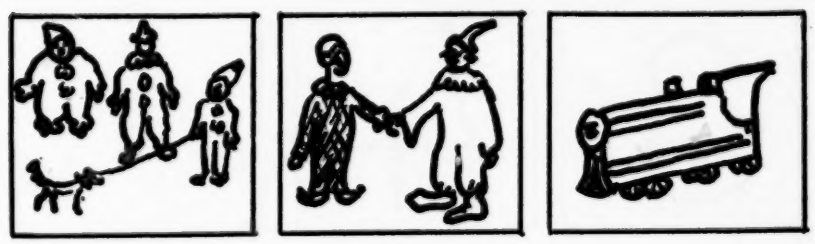


14(6)



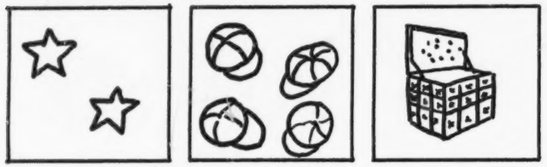
17(7)

15(6)

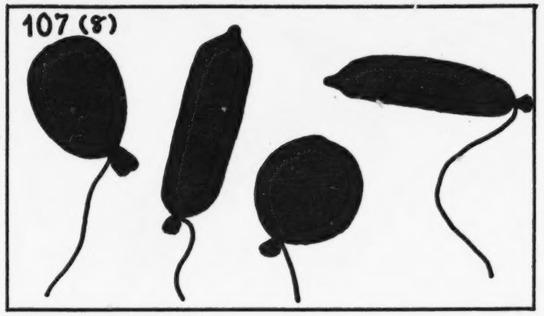


18(2)

19 (3)



107 (8)

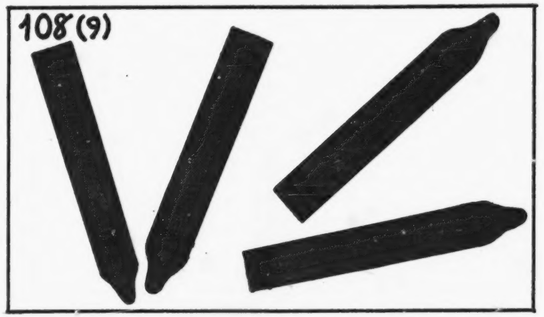


4	2	3

105 (8)

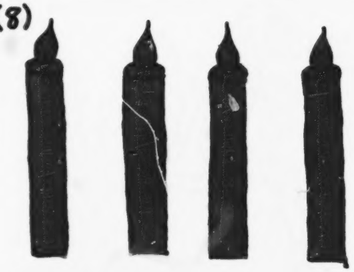
2	4	3

108 (9)



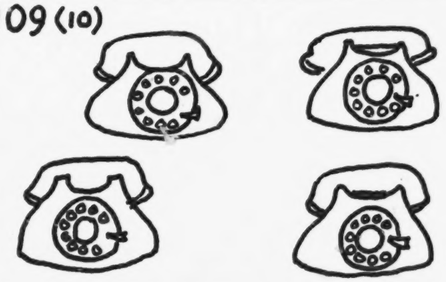
4	3	2

106 (8)



1	3	4

109 (10)



3	4	1

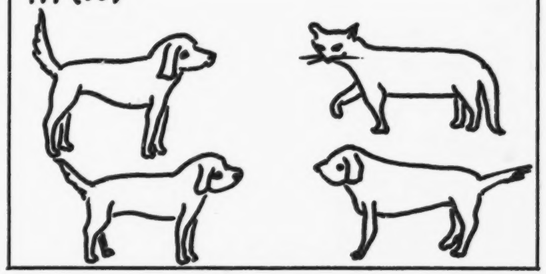
110 (14)

2	3	4

113 (12)

4

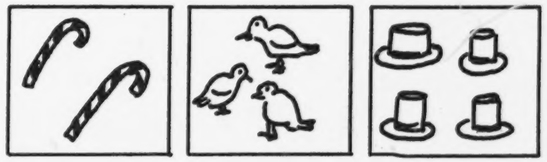
111 (10)



3	2	4

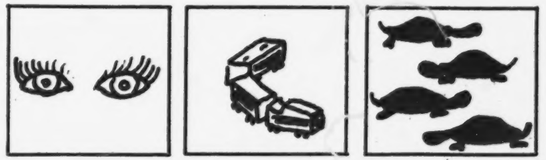
114 (13)

3



112 (11)

4



115 (13)

2

116(12)

4